

¿Matrices Policóricas/Tetracóricas o Matrices Pearson? Un estudio metodológico

Dominguez Lara, Sergio Alexis * a

Artículo Metodológico

ISSN 1852-4206

Abril 2014, Vol. 6, N°1,
39-48

revistas.unc.edu.ar/index.
php/racc

Resumen

El uso de la correlación producto-momento de Pearson es frecuente en la mayoría de estudios en análisis factorial en psicología, mas es conocido que este estadístico solo es aplicable cuando las variables que se relacionan están al menos en escala de intervalo y distribuidas normalmente; su utilización en datos ordinales puede resultar en una matriz de correlaciones distorsionada. Es así que se plantea como una opción adecuada el uso de matrices policóricas/tetracóricas a nivel de ítem en análisis factorial cuando los reactivos están en un nivel de medición nominal u ordinal. El objetivo de este trabajo fue mostrar las diferencias en cuanto a las medidas de adecuación muestral (KMO, test de esfericidad de Bartlett y determinante de la matriz), porcentaje de varianza explicada y saturaciones factoriales tanto en la escala Depresión Rasgo del Inventario de Depresión Rasgo-Estado (Spielberger, Agudelo, y Buela Casal, 2008) como en la dimensión Neuroticismo de la forma corta del Cuestionario de Personalidad de Eysenck-Revisado (Eysenck y Eysenck, 2008) con relación al uso de matrices Policóricas/Tetracóricas y matrices Pearson. Se analizaron dichos instrumentos con diversos métodos de extracción (Máxima verosimilitud, *Minimum Rank Factor Analysis*, Mínimos Cuadrados No Ponderados y Componentes Principales), manteniendo constante el método de rotación Promin. Se apreciaron diferencias con relación a las medidas de adecuación muestral, así como con respecto a la varianza explicada y las cargas factoriales, a favor de las soluciones que tienen como base matrices policóricas/tetracóricas. Se concluye que las matrices policóricas/tetracóricas permiten obtener mejores resultados que las matrices Pearson cuando se trata de análisis factorial a nivel de ítem utilizando distintos métodos.

Palabras claves:

Análisis Factorial; Matrices Tetracóricas; Matrices Policóricas, Validez.

Abstract

Polychoric/Tetrachoric Matrix or Pearson Matrix? A methodological study: The use of product-moment correlation of Pearson is common in most studies in factor analysis in psychology, but it is known that this statistic is only applicable when the variables related are in interval scale and normally distributed, and when are used in ordinal data may to produce a distorted correlation matrix. Thus is a suitable option using polychoric/tetrachoric matrices in item-level factor analysis when the items are in level measurement nominal or ordinal. The aim of this study was to show the differences in the KMO, Bartlett's Test and Determinant of the Matrix, percentage of variance explained and factor loadings in depression trait scale of Depression Inventory Trait - State and the Neuroticism dimension of the short form of the Eysenck Personality Questionnaire -Revised, regarding the use of matrices polychoric/tetrachoric matrices and Pearson. These instruments was analyzed with different extraction methods (Maximum Likelihood, Minimum Rank Factor Analysis, Unweighted Least Squares and Principal Components), keeping constant the rotation method Promin were analyzed. Were observed differences regarding sample adequacy measures, as well as with respect to the explained variance and the factor loadings, for solutions having as polychoric/tetrachoric matrix. So it can be concluded that the polychoric / tetrachoric matrix give better results than Pearson matrices when it comes to item-level factor analysis using different methods.

Key Words:

Factor analysis; polychoric matrix; tetrachoric matrix, validity.

Tabla de Contenido

Introducción	39
Método	41
Participantes	41
Instrumentos	41
Procedimiento	42
Resultados	42
Descriptivos	42
Análisis de Items	42
Análisis Factorial Exploratorio	43
Discusión	44
Referencias	46

Recibido el 15 de Enero de 2014; Recibido la revisión el 24 de Febrero de 2014; Aceptado el 5 de Marzo de 2014.

1. Introducción

El análisis factorial (AF) hace referencia a un conjunto de métodos multivariados de interdependencia que están orientados a identificar una estructura subyacente dentro de un conjunto amplio de datos (Pérez y Medrano, 2010), siendo los

usos más comunes en psicología la evaluación de la estructura interna de un test a partir de los puntajes de los ítems o la comprobación de hipótesis de tipo dimensional usando como medidas puntajes de distintas pruebas (Ferrando y Anguiano-Carrasco,

^a Instituto de Investigación de la Facultad de Psicología y Trabajo Social, Universidad Inca Garcilaso de la Vega.

^{*}Enviar correspondencia a: Dominguez Lara, S.A. E-mail: sdominguezmpcs@gmail.com

2010).

Asimismo, es utilizado para evaluar la unidimensionalidad de un conjunto de reactivos, en los cuales se asume que las respuestas observadas en los ítems constituyen medidas imperfectas de una sola variable latente (Ferrando y Lorenzo-Seva, 1994). En tal sentido, el AF es usado, con frecuencia, para explicar un conjunto de j variables observadas mediante un conjunto pequeño de k variables latentes, ante lo que se espera que los k constructos latentes expliquen la mayor proporción de varianza de la matriz de asociación original $j \times j$, de modo tal que los k variables latentes sirvan para representar las j variables observadas (Henson y Roberts, 2006).

No obstante, debe recalarse que, al observar la ecuación estructural del AF en el caso de un factor común, las variables observables y las latentes son continuas y con distribución normal (Ferrando y Lorenzo-Seva, 1994):

$$z_{im} = a_{mc} * f_{ic} + u_{me} * e_i$$

Donde:

z_{im} = puntuación típica en la variable i

a_{me} = saturación de la variable m en el factor común

f_{ic} = puntuación del sujeto de la variable i en el factor común

u_{me} = saturación de la variable m en el factor específico

e_i = puntuación del sujeto de la variable i en el factor específico

Esto representa un problema, ya que si se pretende utilizar el AF a nivel de ítems, se requiere que dichos reactivos estén al menos en escala de intervalo (Richaud, 2005), lo cual no sucede con muchos de los casos, ya que generalmente se usan formatos multipuntos (escalas de respuesta categórica ordenada o tipo Likert) de tipo ordinal, y los datos a nivel de ítem casi nunca son variables continuas (Ferrando y Anguiano-Carrasco, 2010), siendo en este sentido específicamente problemáticos los dicotómicos u ordinales (Bernstein y Teng, 1989; Ferrando y Anguiano-Carrasco, 2010), dado que un AF de ítems dicotómicos suele dar como resultado una estructura factorial que depende más del grado de dificultad de los ítems (Gaviria, 1990; McDonald y Alhawat, 1974), que del constructo subyacente a los datos (Brown, 2006; McDonald, 1999). Es así que el modelo de AF resulta teóricamente inadecuado (Ferrando y Lorenzo-Seva, 1994).

Esto no representaría dificultad alguna con

variables pertenecientes a otros dominios del conocimiento donde sí se cumplen estas condiciones, pero como señalan Ferrando y Anguiano-Carrasco (2010), los problemas metodológicos del AF en psicología surgen en la primera etapa de la investigación, cuando se planifica el diseño, y siendo una de las decisiones iniciales la del uso de una matriz de correlaciones adecuada y acorde con la naturaleza de los datos, sobre la cual se realizarán los siguientes cálculos implicados en todo el procedimiento.

Contrariamente a la evidencia presentada, el uso de la correlación producto-momento de Pearson es frecuente en la mayoría de estudios en análisis factorial en psicología, al punto que no existe otra referencia del uso de otro tipo de matrices en estudios documentales (Henson y Roberts, 2006) ni en estudios de revisión de buenas prácticas en análisis factorial (Costello y Osborne, 2005; Kahn, 2006; Mvududu y Sink, 2013), mas es conocido que este estadístico solo es aplicable cuando las variables que se relacionan están al menos en escala de intervalo y distribuidas normalmente, y que al ser usadas en datos ordinales puede resultar en una matriz de correlaciones distorsionada (Elosua y Zumbo, 2008).

Hasta el momento queda claro que para que sea adecuado el uso del AF los ítems deben estar en escala de intervalo, pero ¿qué se puede hacer si la propia naturaleza de los ítems no lo permite? Y ¿cómo pueden analizarse datos en escala nominal u ordinal si los fundamentos no lo permiten? Nunnally (1987) plantea que la solución, al margen de las limitaciones teóricas, es tratar a las variables como continuas y proceder con el análisis, utilizando en consecuencia la correlación producto-momento de Pearson.

Otra solución que puede producir una evaluación dimensionalmente adecuada es el uso de las correlaciones policóricas/tetracóricas, la cual se basa en el siguiente postulado: "...sean dos variables, originalmente continuas y distribuidas en forma normal bivalente, que se dicotomizan arbitrariamente. La correlación tetracórica obtenida sobre las dicotomías resultantes es el estimador máximo verosímil de la correlación de Pearson que había entre los continuos originales..." (Ferrando, 1996, p. 401). Es decir, se basan en el supuesto que las categorías de respuesta usadas, sean dicotómica o politómicas, son estimaciones de variables no observables y distribuidas normalmente (Flora y Curran, 2004; O'Connor, 2013).

Por lo tanto, las correlaciones tetracóricas asumen que tras esas dicotomías se encuentran dos variables distribuidas normalmente (Guilford y Fruchter, 1984), así como las correlaciones policóricas las cuales suponen que se trata de variables de intervalo que presentan una distribución normal (Burga, 2006).

Es así que diversos investigadores plantean como una opción adecuada el uso de matrices policóricas/tetracóricas a nivel de ítem en AF de tipo lineal cuando los reactivos están en nivel de medición nominal u ordinal (Elosua y Zumbo, 2008; Ferrando, 1996; Ferrando y Lorenzo-Seva, 1994; O'Connor, 2013; Richaud, 2005), aunque otros no lo recomiendan argumentando que no equivale algebraicamente a la correlación producto-momento de Pearson (Morales-Vallejo, 2013).

En un estudio antecedente a este trabajo, Burga (2006) señala la importancia del AF como uno de los métodos para evaluar la dimensionalidad de un instrumento de medición, considerando problemático el uso de las matrices de correlaciones de Pearson en los AF a nivel de ítems. Recalca que dichas correlaciones no son adecuadas porque los ítems no cumplen con los supuestos necesarios para aplicar la correlación de Pearson: nivel de medición de intervalo y distribución normal de la variable. En tal sentido, rescata el uso de las correlaciones tetracóricas y policóricas, encontrando que la varianza explicada usando matrices policóricas y tetracóricas es mayor que la encontrada utilizando matrices Pearson. La principal limitación de este trabajo es que para considerar la unidimensionalidad de las escalas propuestas solo se utilizó el criterio de varianza compartida, el cual no es concluyente, existiendo otros métodos que dan mejores resultados para la evaluación de unidimensionalidad (Ferrando, 1996).

Cupani y Saurina (2012) estudiaron las propiedades psicométricas del *Self-Directed Search* (Forma E) en una muestra de 340 estudiantes argentinos de colegios públicos y privados. Como parte de su estudio se evaluó la pertinencia del uso de matrices Pearson (ϕ) o matrices tetracóricas, reportando una mayor magnitud de cargas factoriales en las soluciones con matriz de correlaciones tetracóricas. No obstante, decidieron conservar las matrices Pearson (ϕ) dado que no hubo ítems con cargas compartidas en otros factores, y aunque las cargas factoriales hayan sido menores que las

obtenidas con matrices tetracóricas, aún continuaban siendo satisfactorias.

Considerando lo revisado anteriormente, el objetivo de este trabajo fue mostrar las diferencias en cuanto a las medidas de adecuación muestral (índice de adecuación muestral de Kaiser-Meyer-Olkin KMO, test de esfericidad de Bartlett y determinante de la matriz), porcentaje de varianza explicada y saturaciones factoriales tanto en la escala *Depresión Rasgo* del Inventario de Depresión Rasgo-Estado (Spielberger, Agudelo y Buena-Casal, 2008) como en la dimensión Neuroticismo de la forma corta del Cuestionario de Personalidad de Eysenck-Revisado (Eysenck y Eysenck, 2008), utilizando matrices Policóricas/Tetracóricas y matrices Pearson.

2. Método

Se trata de un estudio instrumental, destinado al estudio de las propiedades psicométricas de un test (Montero y León, 2007).

2.1. Participantes

Para la evaluación con el Inventario de Depresión Rasgo-Estado (IDER; Spielberger et al., 2008), la muestra de estudio estuvo conformada por 232 estudiantes de psicología de una universidad privada, siendo 68 varones y 164 mujeres, de edades comprendidas entre 15 y 44 años ($M = 20.89$, $DE = 3.63$). Del total, 221 (95.26%) de los estudiantes encuestados tenían más de 18 años al momento de la evaluación. El muestreo utilizado fue de tipo no aleatorio intencional.

En cuanto a la forma corta del Cuestionario de Personalidad de Eysenck-Revisado (EPQ-R; Eysenck y Eysenck, 2008), la muestra de estudio estuvo conformada por 309 estudiantes de psicología de una universidad privada, 66 varones y 236 mujeres (siete estudiantes no consignaron ese dato). Las edades estuvieron comprendidas entre 16 y 42 años ($M = 20.19$, $DE = 3.75$). Del total, 246 (79.61%) de los estudiantes encuestados tenían más de 18 años al momento de la evaluación. El muestreo utilizado fue no aleatorio de tipo intencional.

2.2. Instrumentos

El Inventario de Depresión Rasgo-Estado (Spielberger et al., 2008), validado en población universitaria (Dominguez, 2011) consta de 20 ítems con formato de respuesta categórica ordenada (tipo Likert) de cuatro puntos. Cuenta con 10 ítems para Depresión-Estado y 10 ítems para Depresión-Rasgo.

Cada escala es independiente, y presente dos factores diferenciados: Eutimia y Distimia. Se utilizó la escala Depresión-Rasgo, que presenta dos subescalas: Eutimia y Distimia.

La forma corta del Cuestionario de Personalidad de Eysenck-Revisado (EPQ-R) propuesta por Eysenck y Eysenck (2008), estudiada en población universitaria (Dominguez, Villegas, Yauri, Aravena, y Ramirez, 2013). Consta de 48 ítems en formato de respuesta dicotómico, con 12 ítems para cada una de las dimensiones que evalúa (Extraversión, Neuroticismo, Psicoticismo y Sinceridad). Se utilizará para este estudio la dimensión Neuroticismo.

2.3. Procedimiento

El inventario se administró como parte de una batería de evaluación psicológica dentro de la casa de estudios. En cuanto a las consideraciones éticas, se les explicó a los estudiantes los objetivos de la evaluación y se les garantizó la confidencialidad de los resultados, los cuales estarían a su disposición luego del procesamiento respectivo, logrando un asentimiento informado por parte de los estudiantes.

Los responsables del estudio explicaron las instrucciones que figuran en el protocolo de aplicación de la prueba, y despejaron las dudas que las personas evaluadas tuvieron. Posteriormente, se elaboró una base de datos en MS Excel.

El análisis de los datos se llevó a cabo utilizando el programa ViSta (Young, 2003) para el análisis de ítems, y posteriormente se usó FACTOR (Lorenzo-Seva y Ferrando, 2007) para la realización del AF exploratorio.

Se trabajaron los siguientes aspectos en la mencionada base de datos: Método de extracción de factores: Máxima verosimilitud / Minimum Rank Factor Analysis / Mínimos Cuadrados No Ponderados / Componentes Principales. Método de Rotación: Promin (se mantuvo constante).

Cabe mencionar que no fue posible realizar el método de Máxima Verosimilitud tanto para matrices policóricas como tetracóricas en el IDER, y en el caso de los ítems dicotómicos del EPQ-R, no se llegó a una convergencia que permita el procedimiento de Máxima Verosimilitud para matrices Pearson.

3. Resultados

3.1. Estadísticos descriptivos

Se analizaron los estadísticos descriptivos de cada uno de los reactivos del instrumento, los cuales se

muestran en la [Tabla 1](#). En la mayoría de reactivos se aprecian indicadores descriptivos dentro de rangos aceptables, aunque en algunos reactivos existen elevados indicadores de asimetría y curtosis.

3.2. Análisis de ítems

En ambos instrumentos estudiados, todos los reactivos presentan una adecuada correlación ítem-test. Esto es importante debido a que, más allá de la verificación de los índices de discriminación, se pueden descartar aquellos reactivos con baja confiabilidad y eliminarlos, a fin de realizar el AF con menos ruido (Ferrando y Anguiano-Carrasco, 2010).

Tabla 1. Estadísticos descriptivos del Inventario de Depresión Estado/Rasgo-R.

	M	DE	As	Cs	Correlación ítem-test corregida*
Ítem 1	3.341	.771	-.898	-.025	.654
Ítem 2	1.341	.602	1.942	4.116	.315
Ítem 3	2.703	.842	-.045	-.687	.453
Ítem 4	2.935	.831	-.196	-.868	.498
Ítem 5	3.444	.729	-1.112	.479	.555
Ítem 6	1.517	.657	1.177	1.340	.519
Ítem 7	1.638	.718	1.087	1.212	.380
Ítem 8	1.431	.657	1.152	.963	.531
Ítem 9	1.547	.614	0.772	.182	.469
Ítem 10	2.978	.774	-.355	-.351	.586

Nota: M = Media; DE = Desviación Estandar; As = Asimetría; Cs = Curtosis. *Realizado con los ítems 2, 6, 7, 8 y 9 en escala invertida

Con relación al análisis de ítems de la dimensión Neuroticismo del EPQ-R, la mayoría de los ítems se encuentran entre .40 y .60, rango considerado como adecuado según Ferrando y Anguiano-Carrasco (2010), el cual minimiza la posibilidad de una relación no lineal entre las variables ([Tabla 2](#)).

Tabla 2. Estadísticos descriptivos del EPQ-R (N).

	M	Ds	As	Cs	Correlación ítem-test corregida	Índice de Dificultad
Ítem 2	.350	.476	0.633	-1.597	.548	.739
Ítem 4	.217	.412	1.379	-.102	.522	.511
Ítem 8	.246	.430	1.184	-.600	.365	.434
Ítem 13	.155	.362	1.909	1.636	.556	.422
Ítem 18	.243	.429	1.204	-.552	.563	.591
Ítem 19	.489	.500	.045	-1.995	.464	.663
Ítem 20	.159	.365	1.875	1.509	.496	.403
Ítem 24	.340	.473	.679	-1.538	.568	.725
Ítem 32	.246	.430	1.184	-.600	.571	.598
Ítem 35	.450	.497	.202	-1.956	.330	.489
Ítem 41	2.33	.423	1.267	-.396	.547	.549
Ítem 42	2.33	.423	1.1267	-.396	.528	.553

Nota: M = Media; DE = Desviación Estandar; As = Asimetría; Cs = Curtosis.

Tabla 3. Cargas factoriales del IDER-R según método de extracción factorial y matriz de correlaciones.

	MV						MRFA					
	M. Pearson			M. Policórica			M. Pearson			M. Policórica		
	MC	F1	F2	ME	F1	F2	MC	F1	F2	ME	F1	F2
Item 2	.476	.015	.470	-.194	.524	.507	.731	.015	.723	-.363	.723	-.363
Item 6	.709	-.030	.722	-.341	.755	.762	.779	-.033	.796	-.436	.796	-.436
Item 7	.613	.049	.592	-.220	.589	.575	.676	.042	.654	-.308	.654	-.308
Item 8	.605	-.121	.638	-.387	.682	.726	.752	-.093	.800	-.483	.800	-.483
Item 9	.736	.036	.720	-.288	.723	.714	.734	-.005	.737	-.386	.737	-.386
Item 1	-.040	.792	-.388	.810	-.046	.797	-.084	.782	-.489	.825	-.489	.825
Item 3	.143	.734	-.179	.671	.150	.753	.210	.763	-.186	.654	-.186	.654
Item 4	.100	.743	-.226	.699	.100	.743	.113	.747	-.274	.689	-.274	.689
Item 5	-.112	.579	-.366	.628	-.113	.591	-.204	.593	-.511	.699	-.511	.699
Item 10	-.157	.583	-.413	.652	-.147	.609	-.108	.591	-.414	.647	-.414	.647
Determinante de la Matriz	.044											
Test de Esfericidad de Bartlett	707.5; p < .001											
KMO	.847											
Autovvalor Inicial	3.821											
% de VE	.5589											
% de VCE	.8307											
Corr. entre Fact.	-.422											

Nota. MV: Máxima Verosimilitud; MRFA: Minimum Rank Factor Analysis; MC: Matriz de Configuración; ME: Matriz de Estructura; F1: Factor 1; F2: Factor 2; VE: Varianza Explicada; VCE: Varianza Común Explicada; Corr. entre Fact.: Correlación entre factores

Tabla 3 (continuación). Cargas factoriales del IDER-R según método de extracción factorial y matriz de correlaciones.

	MCNP						CP									
	M. Pearson			M. Policórica			M. Pearson			M. Policórica						
	MC	F1	F2	ME	F1	F2	MC	F1	F2	ME	F1	F2				
Item 2	.480	.022	.470	-.188	.688	-.004	.690	-.372	.623	.068	.600	-.151	.768	.009	.764	-.329
Item 6	.719	-.026	.730	-.341	.763	-.034	.781	-.442	.766	-.048	.783	-.318	.794	-.055	.818	-.405
Item 7	.613	.049	.591	-.220	.699	.061	.666	-.313	.726	.072	.701	-.184	.789	.084	.752	-.264
Item 8	.609	-.118	.661	-.385	.692	-.112	.752	-.482	.678	-.137	.727	-.376	.736	-.128	.792	-.453
Item 9	.719	.025	.708	-.291	.734	.001	.733	-.392	.771	.006	.769	-.265	.787	-.007	.790	-.354
Item 1	-.046	.788	-.392	.809	-.085	.770	-.497	.815	-.075	.804	-.358	.831	-.133	.771	-.472	.829
Item 3	.143	.731	-.178	.668	.199	.739	-.197	.633	.168	.819	-.120	.760	.240	.852	-.136	.746
Item 4	.109	.750	-.220	.703	.125	.758	-.281	.691	.121	.824	-.169	.782	.133	.837	-.236	.778
Item 5	-.114	.581	-.369	.631	-.205	.580	-.515	.689	-.117	.669	-.352	.711	-.231	.633	-.510	.735
Item 10	-.156	.582	-.412	.651	-.117	.567	-.421	.630	-.165	.661	-.398	.719	-.122	.652	-.410	.706
Determinante de la Matriz	.044															
Test de Esfericidad de Bartlett	707.5; p < .001															
KMO	.847															
Autovvalor Inicial	3.821															
% de VE	.5589															
Corr. entre Fact.	-.439															

Nota. MCNP: Mínimos Cuadrados No Ponderados; CP: Componentes Principales; MC: Matriz de Configuración; ME: Matriz de Estructura; F1: Factor 1; F2: Factor 2; C1: Componente 1; C2: Componente 2; VE: Varianza Explicada; Corr. entre Fact.: Correlación entre factores

3.3. Análisis Factorial Exploratorio

En los dos instrumentos citados anteriormente, para cada uno de los análisis propuestos (Máxima verosimilitud/Análisis Factorial de Rango Mínimo/Mínimos Cuadrados No Ponderados/Componentes Principales), se exploraron las medidas de adecuación muestral, es decir, si los ítems poseían la suficiente correlación para realizar el análisis factorial (Tablas 3 y 4). Esto se realizó mediante el análisis de la determinante de la matriz, el KMO y el test de Esfericidad de Bartlett. Luego se realizaron los análisis planificados anteriormente, usando tanto matrices policóricas/tetracóricas y

matrices Pearson. Mediante el método clásico de análisis paralelo (Horn, 1965) como con el método óptimo (Timmerman y Lorenzo-Seva, 2011) se sugirió la extracción de dos factores. Se trabajó además con la rotación Promin (Lorenzo-Seva, 1999, 2013), dado que es el método de rotación oblicua más recomendado para AF cuando se aprecia una correlación por encima de .30 entre los factores y, además, cuando la teoría sustenta dicha correlación (Ferrando y Anguiano-Carrasco, 2010), como es el caso de los dos factores hallados en el IDER: Eutimia y Distimia.

De acuerdo con los resultados de los tres

métodos de AF estudiados en cada instrumento (Máxima verosimilitud/Minimum Rank Factor Analysis/Mínimos Cuadrados No Ponderados), así como el referido a reducción de variables (Componentes Principales), se hallaron indicadores adecuados previos al AF que permitan continuar el procedimiento (determinante de la matriz de correlaciones significativa, índice de adecuación muestral KMO adecuado y test de esfericidad de Bartlett significativo). No obstante, se aprecian algunas diferencias con relación a estos últimos, así como con respecto a la varianza explicada y las cargas factoriales según el uso de matrices policóricas y tetracóricas con relación al uso de matrices Pearson.

Respecto a ello, en las soluciones que implicaron una matriz de correlaciones producto-momento de Pearson la determinante de la matriz fue cercana a .05, siendo aquella relacionada con las matrices de correlaciones policóricas menor que dicho número, lo cual indica que la intercorrelación entre los ítems evaluada mediante correlaciones producto-momento de Pearson es elevada, pero no lo suficiente; del mismo modo, el estadístico asociado al test de esfericidad de Bartlett y el índice de adecuación muestral KMO fueron menores en las matrices Pearson.

4. Discusión

El objetivo de este trabajo fue evaluar la pertinencia del uso de las matrices policóricas/tetracóricas con respecto a las matrices Pearson, destacando las diferencias existentes en diversos indicadores relacionados con el análisis factorial según el método de extracción de factores empleado.

Los análisis realizados de forma previa sobre el IDER (Agudelo, 2009; Agudelo, Spielberger, y Buela-Casal, 2007; Dominguez, 2011; Ocampo, 2007; Vera et al., 2008) dejó claro que los dos factores en cada una de las escalas (Estado y Rasgo), eutimia y distimia, son elementos diferenciados, tanto teórica como empíricamente. Del mismo modo, el uso de la dimensión Neuroticismo del EPQ-R fue elegida debido a que es la escala más confiable en los estudios realizados con dicho instrumento en diversos contextos (Allik, 2005; Dominguez et al., 2013; Francis, Alan, y Ziebertz, 2006; Francis, Brown, y Philipchalk, 1992; Ibañez, 1997; Katz y Francis, 2000; Schmidt et. al., 2008; Tiwari, Singh, y Singh, 2009). Esto

posibilitó un análisis más objetivo, ya que toda diferencia observada en cuanto a los indicadores, haría referencia al uso de las matrices policóricas/tetracóricas frente al uso de las matrices Pearson.

Se realizó un análisis de ítems a ambos instrumentos, y a partir de ello se apreció que todos los reactivos presentaban índices de discriminación adecuados, así como índices de dificultad dentro del rango esperado (Aiken, 2003) para el caso de la dimensión Neuroticismo del EPQ-R, lo cual indica que no se podría esperar grandes distorsiones en solución factorial (Cupani y Saurina, 2012), ni factores de dificultad (Gaviria, 1990).

En todas las soluciones, si bien es cierto que las diferencias con relación a las medidas de adecuación muestral no son marcadas, podrían acentuarse en instrumentos que sean de reciente creación o en procesos de adaptación debido a que estas dependen estrechamente de la matriz de correlaciones utilizada y de la población en la cual se esté utilizando, lo cual podría brindar un indicador que infravalore la capacidad de intercorrelación de los ítems, dado que un test de Bartlett no significativo o un KMO debajo del nivel esperado, podría indicar que las variables no están lo suficientemente intercorrelacionadas, cuando realmente sí lo están.

Por su parte, el porcentaje de la varianza explicada por aquellas soluciones que utilizaron matrices policóricas o tetracóricas fue mayor, apreciándose con relación a esta última las mayores diferencias. El porcentaje de varianza explicada en cada análisis realizado en el IDER indicaría unidimensionalidad (Carmines y Zeller, 1979), aunque reportes más recientes abogan por la integración de la información de diversos indicadores que brinden tal evidencia (Ferrando, 1996). Es conveniente mencionar que la varianza explicada producto de la matriz de correlaciones policóricas fue mayor. Asimismo, en el caso de las soluciones a partir de matrices Pearson, en la dimensión Neuroticismo del EPQ-R el porcentaje de varianza explicada no llegó a 40%, siendo el porcentaje de la varianza explicada a partir del uso de matrices tetracóricas y Pearson mucho mayor.

Con relación a las soluciones factoriales, los coeficientes estructurales (cargas factoriales) de las matrices de estructura en aquellas soluciones que tienen como base las matrices policóricas, y las cargas

factoriales de las matrices sin rotar de las soluciones resultantes a partir de las matrices tetracóricas, en su mayoría fueron de mayor magnitud que sus contrapartes relacionadas con las matrices Pearson. Además de ello, las cargas factoriales en todos los casos fueron superiores a .40 (Glutting, 2002; Zwick y

Velice, 1986), lo cual además de ser adecuadas, sería un indicador de relevancia práctica y de una correcta definición del factor (Hair, Anderson, Tatham, y Black, 1999).

Tabla 4.

Cargas factoriales del EPQ-R (N) según método de extracción factorial y matriz de correlaciones.

	MRFA		MCNP		CP	
	M. Pearson	M. Policórica	M. Pearson	M. Policórica	M. Pearson	M. Policórica
Ítem 2	.636	.758	.619	.740	.662	.766
Ítem 4	.584	.728	.578	.723	.626	.751
Ítem 8	.399	.536	.397	.527	.451	.571
Ítem 13	.637	.802	.608	.774	.653	.794
Ítem 18	.626	.776	.621	.764	.665	.786
Ítem 19	.535	.657	.509	.661	.563	.652
Ítem 20	.538	.704	.543	.709	.595	.739
Ítem 24	.679	.770	.620	.727	.665	.756
Ítem 32	.644	.782	.636	.779	.678	.798
Ítem 35	.354	.443	.359	.444	.411	.488
Ítem 41	.620	.772	.589	.734	.637	.762
Ítem 42	.637	.756	.596	.729	.642	.755
Determinante de la Matriz	.023	.0001	.023	.0001	.023	.0001
Test de Esfericidad de Bartlett	1137.9; p<.001	2741.3; p<.001	1137.9; p<.001	2741.3; p<.001	1137.9; p<.001	2741.3; p<.001
KMO	.860	.818	.860	.818	.860	.818
Autovalor Inicial	4.464	6.296	4.464	6.296	4.464	6.296
% de VE	.3720	.5247	.3720	.5247	.3720	.5247
% de VCE	.6008	.6409	-	-	-	-

Nota. MRFA: Minimum Rank Factor Analysis; MCNP: Mínimos Cuadrados No Ponderados; CP: Componentes Principales; VE: Varianza Explicada; VCE: Varianza Común Explicada.

En este último aspecto cabe resaltar que fueron diversos métodos de extracción de factores que fueron utilizados para cumplir con los objetivos de este trabajo, aunque el procedimiento de máxima verosimilitud (MV) no admitió el uso de matrices policóricas/tetracóricas para su estimación. Esto se da debido a que el método de MV se basa en el supuesto que las variables son continuas, métricas y con distribución normal multivariante, aspectos que no se asumen en las mencionadas matrices. En cambio, el método de mínimos cuadrados no ponderados (MCNP) no se basa en esos supuestos dado que tiene carácter descriptivo, y puede usarse cuando se viola el supuesto de normalidad, aunque en la práctica puede tener similitudes con la solución factorial producto del método de MV, siendo uno de los métodos recomendados (Ferrando y Anguiano-Carrasco, 2010).

Del mismo modo, se trabajó con el Minimum Rank Factor Analysis (MRFA; Shapiro y Ten Berge, 2002), debido a que es capaz de distinguir la varianza

común explicada de la varianza total común, y evaluarla por separado, y es capaz de detectar aquellas soluciones inadecuadas. En los resultados obtenidos, además de mejores cargas factoriales, se aprecia una mayor parte de varianza común explicada en aquellas soluciones que parten de matrices policóricas y tetracóricas, con respecto a las soluciones a partir de matrices Pearson.

Por último, se trabajó con el análisis de componentes principales (ACP), ya que si bien no es una técnica de análisis factorial propiamente dicha, sino más bien una técnica de reducción de información (Ferrando y Anguiano-Carrasco, 2010; Henson y Roberts, 2006; Pérez y Medrano, 2010), puede ser usada cuando se desea estimar la mayor parte de la varianza encontrada en un conjunto de datos (Elosua y Zumbo, 2008).

En el presente análisis, tanto en lo concerniente a matrices policóricas como tetracóricas, el ACP es el método que presenta las saturaciones más elevadas

con relación a los otros métodos (MV, MRFA y MCNP). Esto puede explicarse debido a que el ACP al no ser una técnica de AF, no distingue entre varianza común y específica, por lo que tiende a inflar los valores hallados en la matriz de componentes (Costello y Osborne, 2005), sobre todo si hay pocos reactivos (Morales-Vallejo, 2013). En tal sentido, se consideraría en este caso un incremento debido a las causas mencionadas.

Entonces, a partir de los hallazgos presentados, se puede concluir que las matrices policóricas/tetracóricas dan mejores resultados que las matrices Pearson cuando se trata de AF utilizando distintos métodos, lo cual está en consonancia con algunos estudios similares realizados anteriormente (Burga, 2006; Cupani y Saurina, 2012).

Entre las limitaciones se destaca que en este estudio se analizan dos instrumentos que fueron elegidos en función de características particulares en relación a estudios previos y tipo de respuesta (uno con formato de respuesta politómico y otro dicotómico), lo cual no los hace representativos de todos los instrumentos de evaluación psicológica, mas brinda alcances respecto a las consideraciones que se deben tener al momento de planificar un AF.

Del mismo modo, la muestra está sesgada hacia un grupo en particular (estudiantes universitarios) el cual fue tomado en consideración por el acceso que se tenía del mismo y no por razones de índole teórica ni metodológica. Ante ello, sería recomendable realizar estudios similares con población general y con una ampliación del tipo de instrumentos a fin de poder replicar y extrapolar las conclusiones presentadas en este trabajo, el cual serviría entonces como un antecedente para continuar investigando la temática abordada.

Para terminar, si bien es cierto que las matrices policóricas/tetracóricas son una buena alternativa en AF de tipo lineal, actualmente hay desarrollos del AF ligados a la teoría de respuesta al ítem (TRI), donde se considerarían relaciones no lineales entre los ítems y el factor común (Ferrando, 1994; Ferrando y Lorenzo-Seva, 2013; Richaud, 2005).

Referencias

- Agudelo, D. (2009). Propiedades psicométricas del Inventario de Depresión Estado/Rasgo (IDER) con adolescentes y universitarios de la ciudad de Bucaramanga. *Pensamiento Psicológico*, 5(12), 139-160.
- Aiken, L. (2003). *Tests psicológicos y evaluación*. México: Prentice Hall.
- Allik, J. (2005). Personality dimensions across cultures. *Journal of Personality Disorders*, 19(3), 212-232.
- Agudelo, D., Spielberger, C., & Buéla-Casal, G. (2007). La depresión: ¿un trastorno dimensional o categorial? *Salud Mental*, 30, 20-28.
- Bernstein, I., & Teng, G. (1989). Factoring ítems and factoring scales are different: Spurious evidence for multidimensionality due to ítem categorization. *Psychological Bulletin*, 105, 467-477.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.
- Burga, A. (2006). La unidimensionalidad de un instrumento de medición: perspectiva factorial. *Revista de Psicología*, 24(1), 53-80
- Carmine, E., & Zeller, R. (1979). *Reliability and validity assessment*. London: Sage.
- Costello, A., & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7), 1-9. Disponible en: <http://pareonline.net/pdf/v10n7.pdf>
- Cupani, M., & Saurina, I. (2012). Estudios psicométricos del Self-Directed Search (Forma E) en una muestra de estudiantes argentinos. *Evaluar*, 11, 1-17.
- Dominguez, S. (2011). *Propiedades Psicométricas del Inventario de Depresión Estado Rasgo en estudiantes universitarios*. VIII Congreso Nacional de Profesionales y Estudiantes de Psicología. Facultad de Psicología y Trabajo Social. Universidad Inca Garcilaso de la Vega. Lima, Perú.
- Dominguez, S., Villegas, G., Yauri, C., Aravena, S., & Ramírez F. (2013). Análisis Psicométrico Preliminar de la Forma Corta del EPQ-R en Estudiantes Universitarios. *Avances en Psicología*, 21(1), 73-82.
- Elosua, P., & Zumbo, B. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema*, 20(4), 896-901.
- Eysenck, H. & Eysenck, S. (2008). *Cuestionario revisado de Personalidad de Eysenck versiones completa (EPQ-R) y abreviada (EPQ-RS)*. Madrid: TEA Ediciones.
- Ferrando, P. (1996). Evaluación de la unidimensionalidad de los ítems mediante análisis factorial. *Psicothema*, 8(2), 397-410.
- Ferrando, P., & Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo*, 31(1), 18-33.
- Ferrando, P., & Lorenzo-Seva, U. (1994). Recuperación de la solución factorial a partir de variables dicotomizadas. *Psicothema*, 6(3), 483-491.
- Ferrando, P. J., & Lorenzo-Seva, U. (2013). *Unrestricted item factor analysis and some relations with item response theory*. Technical Report. Department of Psychology, Universitat Rovira i Virgili, Tarragona. Document

- available at: <http://psico.fcep.urv.cat/utilitats/factor/>
- Flora, D., & Curran, P. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis in psychological research. *Psychological Methods, 9*, 466-491.
- Francis, L., Alan, C., & Ziebertz, H. (2006). The short-form Revised Eysenck Personality Questionnaire (EPQR-S): a German edition. *Social Behavior and Personality, 34*(2), 197-203.
- Francis, L., Brown, L., & Philipchalk, R. (1992). The development of an abbreviated form of the Revised Eysenck Personality Questionnaire (EPQR-A): its use among students in England, Canada, the U.S.A. and Australia. *Personality and Individual Differences, 13*(4), 443-449.
- Gaviria, J. (1990). Factores de dificultad en el análisis de ítems. Qué son, por qué aparecen y posibles soluciones. *Revista Complutense de Educación, 1*(1), 95-108.
- Glutting, J. (2002). Some psychometric properties of a system to measure ADHD among college students. *Measurement and Evaluation in Counseling and Development, 34*, 194-209.
- Guilford, P., & Fruchter, B. (1984). *Estadística aplicada a la psicología y la educación*. México: McGraw-Hill.
- Hair, J., Anderson, R., Tatham, R., & Black, W. (1999). *Análisis multivariante* (5ª ed.). Madrid: Prentice Hall.
- Henson, R., & Roberts, J. (2006). Use of exploratory factor analysis in published research. Common errors and some comment on improved practice. *Educational and Psychological Measurement, 66*(3), 393-416.
- Horn, J. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.
- Ibañez, I. (1997). *Primeros datos de la versión en castellano del Cuestionario Revisado de Personalidad de Eysenck (EPQ-R)*. Recuperado el 19 abril, 2012, de: www.uji.es/bin/publ/edicions/jf11/eysenck.pdf
- Kahn, J. (2006). Factor Analysis in Counseling Psychology Research, Training, and Practice: Principles, Advances, and Applications. *The Counseling Psychologist, 34*(5), 684-718.
- Katz, Y., & Francis, L. (2000). Hebrew revised Eysenck Personality Questionnaire: short form (EPQR-S) and abbreviated form (EPQR-A). *Social Behavior and Personality, 28*(6), 555-560.
- Lorenzo-Seva, U. (1999). Promin: a method for oblique factor rotation. *Multivariate Behavioral Research, 34*, 347-356.
- Lorenzo-Seva, U. (2013). *Why rotate my data using Promin?* Technical Report. Department of Psychology, Universitat Rovira i Virgili, Tarragona. Documento disponible en: <http://psico.fcep.urv.cat/utilitats/factor/>
- Lorenzo-Seva, U., & Ferrando, P. (2007). *FACTOR: A computer program to fit the exploratory factor analysis model*. University Rovira i Virgili.
- McDonald, R. (1999). *Test theory: A unified treatment*. Mahwah, NJ: LEA.
- McDonald, R., & Alhawat, K. (1974). Difficulty factors in binary data. *British Journal of Mathematical and Statistical Psychology, 27*, 82-89.
- Montero, O., & León, I. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology, 7*(3), 847-862.
- Morales-Vallejo, P. (2013). *El Análisis Factorial en la construcción e interpretación de tests, escalas y cuestionarios*. Documento disponible en: <http://www.upcomillas.es/personal/peter/investigacion/AnalisisFactorial.pdf>
- Mvududu, N., & Sink, C. (2013). Factor Analysis in Counseling Research and Practice. *Counseling Outcome Research and Evaluation, 4*(2), 75-98.
- Nunnally, J. (1987). *Teoría Psicométrica*. México: Trillas.
- Ocampo, L. (2007). Análisis correlacional del Cuestionario de Depresión Estado/Rasgo (ST-DEP) con una muestra de adolescentes y universitarios de Medellín. *Psicología desde el Caribe, 20*, 28-49.
- O'Connor, B. (2013). *Cautions Regarding Item-Level Factor Analyses*. Recuperado desde: <https://people.ok.ubc.ca/brioconn/nfactors/itemanalysis.html>.
- Pérez, E., & Medrano, L. (2010). Análisis factorial exploratorio: Bases conceptuales y metodológicas. *Revista Argentina de Ciencias del Comportamiento, 2*(1), 58-66.
- Richaud, M. (2005). Desarrollos del análisis factorial para el estudio de ítem dicotómicos y ordinales. *Interdisciplinaria, 22*(2), 237-251.
- Shapiro, A., & Ten Berge, J.M.F. (2002). Statistical inference of minimum rank factor analysis. *Psychometrika, 67*, 79-94.
- Schmidt, V., Firpo, L., Vion, D., De Costa Oliván, D., Casella, L., Cuenya, L., Blum, G., & Pedrón, V. (2008). Modelo psicobiológico de la personalidad de Eysenck: una historia proyectada hacia el futuro. *Revista Internacional de Psicología, 11*(2), 1-21.
- Spielberger, C., Agudelo, D., & Buela-Casal, G. (2008). *Inventario de Depresión Rasgo/Estado (IDER)*. Madrid: TEA Ediciones.
- Timmerman, M., & Lorenzo-Seva, U. (2011). Dimensionality Assessment of Ordered Polytomous Items with Parallel Analysis. *Psychological Methods, 16*(2), 209-220.
- Tiwari, T., Singh, A., & Singh, I. (2009). The short-form Revised Eysenck Personality Questionnaire: a Hindi edition (EPQRS-H). *Industrial Psychiatry Journal, 18*(1), 27-31.
- Young, F. (2003). ViSta "The Visual Statistics System". [programa informático]. Recuperado de: URL: <http://forrest.psych.unc.edu/research/index.html>, el

16.11.11.

Vera, P., Buela-Casal, G., Celis, K., Córdova, N., Encina, N., & Spielberger, C. (2008). Chilean experimental version of the State/Trait depression Questionnaire (ST/DEP): Trait subscale. *International Journal of Clinical and Health Psychology, 8*, 563-575.

Zwick, W., & Velicer, W. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442.